



berns language consulting

blc-Paper

# Maschinelle Übersetzung

## Der Weg zur eigenen MT-Engine



## Sie haben Fragen? Kontaktieren Sie uns!

berns language consulting GmbH

Blücherstraße 69  
40477 Düsseldorf



Zentrale +49 (0) 211 22 06 77 0

[www.berns-language-consulting.de](http://www.berns-language-consulting.de)

[in linkedin.com/company/berns-language-consulting](https://www.linkedin.com/company/berns-language-consulting)

[Besuchen Sie unsere Leistungsseite zu  
Maschinellem Übersetzung](#)



# Wie kommen Sie zu Ihrer eigenen MT-Lösung?

## Einleitung: Übersetzung im Wandel

## MT-Revolution: Von Regeln zu Neuronen

## Use-Case: MT bis der Arzt kommt

## Kontakt: Mit blc zur eigenen MT-Lösung

# Übersetzung im Wandel



Im Zeitalter der **künstlichen Intelligenz** ist alles dauerhaft im Wandel! Übersetzungstechnologie, Content Management, der Umgang mit der Cloud und die kollaborative Zusammenarbeit im Netz. Keine Technologie hat in der letzten Zeit einen so spürbaren Einfluss auf die Übersetzungslandschaft genommen, wie die **maschinelle Übersetzung** (kurz MT für ‚Machine Translation‘). Mit schneller und qualitativ hochwertiger maschineller Übersetzung ändern sich nicht nur die Anforderungen an Übersetzerleistungen, sondern auch die Gestaltung der Übersetzungsworkflows, die Prozesse zur Qualitätssicherung und die Preisgestaltung der Dienstleister.

Um maschinelle Übersetzung zum integralen Bestandteil Ihrer **Übersetzungslandschaft** zu machen, sollten Sie die Potenziale dieser Technologie kennen und vollständig ausschöpfen.

Unser **blc-Paper** liefert nützliche Hintergrundinformationen zur aktuellen maschinellen Übersetzungstechnologie und einen Überblick zu den notwendigen Schritten auf dem Weg zur eigenen maschinellen Übersetzungseingabe.

Durch den Einsatz maschineller Übersetzung werden Übersetzungskosten und Lieferzeiten extrem reduziert. Das volle Potenzial dieser Technologie kann jedoch nur ausgeschöpft werden, wenn alle Aspekte der maschinellen Übersetzung sorgfältig geplant und umgesetzt werden. Hierzu zählen vor allem:

- Extraktion und Aufbereitung der **Trainingsdaten**
- **Auswahl** des Engine-Herstellers
- Training und **Optimierung** der Engines
- Automatische und **humangestützte Evaluation** zur Feststellung des Qualitätsniveaus
- Integration der Engines und **Post-Editing**.

Wir geben Ihnen einen Überblick zu Anforderungen, Bedingungen und Beistellungen, die Ihrem **MT-Vorhaben** zum Erfolg verhelfen.

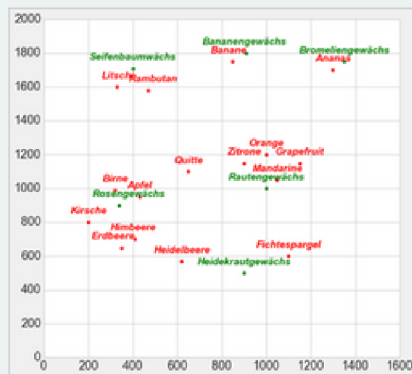
# Von Regeln zu Neuronen

Seit den 60er Jahren versuchen MT-Forscher bereits, grammatikalische und lexikalische Sprachbildung mit Regeln für die Übersetzung nutzbar zu machen. Während regelbasierte Systeme (RBMT-Systeme) schon damals mit Glossaren an kundenspezifische Terminologie angepasst werden konnten, war die Anpassung an den eigenen Stil nur begrenzt möglich. Um die Jahrtausendwende machten die fortschreitende Digitalisierung von Textinhalten, ihre Verbreitung im Internet und steigende Rechnerkapazitäten den Weg frei für eine andere Herangehensweise in der maschinellen Übersetzung: Man entwickelte statistische Systeme (SMT), die auf Basis umfangreicher Textmengen und Übersetzungen wichtige Beziehungen zwischen Ausgangstext und Übersetzung lernten. Von einer menschenähnlichen Übersetzungsqualität war man hier immer noch weit entfernt.

Kaum 15 Jahre später erleben wir eine **Revolution** in der automatischen Sprachverarbeitung: **DeepLearning** – ein auf künstlichen neuronalen Netzen basierendes Lernverfahren – hat sprachbasierte Anwendungen nachhaltig verändert. Eine der ersten Technologien, die vom DeepLearning profitierten, war die neuronale maschinelle Übersetzung (NMT).

Wie bei der statistischen MT ist auch für die neuronale MT ein zweisprachiger Text Voraussetzung für das Training einer Engine. Das Verfahren dahinter beruht auf einer **abstrakten Repräsentation** von Wörtern in einem **mehrdimensionalen Vektorraum** (siehe Grafik). Die Repräsentationen werden in einer komplexen Netzwerkarchitektur verarbeitet, die Gewichtungen für Wörter und Übersetzungen lernt und optimiert.

Im Unterschied zu Large Language Models (LLMs), die Ihre Stärke in der Generierung von einsprachigen Texten haben, ist die NMT auf den Bedeutungstransfer zwischen zwei Sprachen spezialisiert. Eine sehr gute **Lesbarkeit** der neuronalen Übersetzungen geht mit **geringen Nachbearbeitungsaufwänden** einher. Aus diesem Grund ist das NMT-Verfahren für den Produktiveinsatz in Übersetzungsworkflows und als Selbstbedienungsservice sehr attraktiv.



Vektorrepräsentationen als Grundlage für maschinelle Übersetzung

# Datenbasis, Training, Evaluation & Integration

Der Weg zur eigenen MT-Engine wird mit möglichst umfangreichen **Sprachdaten** und durch klar definierte **Prozessschritte** bereitet. Die notwendigen Prozessschritte beinhalten insbesondere die Absicherung der Datenqualität und das präzise Testen der trainierten MT-Engines.

Unabhängig davon, ob MT-Engine-Training zur Evaluation von Systemanbietern durchgeführt wird oder ob der Anbieter bereits feststeht: Der **Weg zur MT-Engine** kann anhand von vier klar definierten Workflow-Schritten beschrieben werden:



## Datenextraktion

Zur **Erstellung** der **Trainingsdatenbasis** werden alle verfügbaren Datenquellen identifiziert und nutzbar gemacht. Translation Memories sowie Daten aus XML-Dateien, Office-Tabellen, Webinhalten und Services werden hierzu herangezogen.

## Datenaufbereitung

Vor dem MT-Training werden die extrahierten Daten auf **linguistische** und **formale Eignung** geprüft. Hierbei ist vor allem die Domäne (Sachgebiet) der zu übersetzenden Texte relevant, sowie datenschutzrechtlich sensible Inhalte, die Gegenstand einer Anonymisierung sein sollen

## Evaluation & Integration

Zur Messung der Qualität von MT-Engines sind automatische Scores, wie z.B. BLEU, nur bedingt geeignet. Um Problemfälle in der MT genau zu identifizieren, werden sprachliche **Bewertungen** der Textinhalte von menschlichen Experten durchgeführt. Gleichzeitig wird die Post-Edit-Distanz laufend ermittelt.

## Training

Die Ergebnisqualität des MT-Trainings ist vor allem abhängig von der Zusammenstellung der Übersetzungseinheiten, von den **Trainingsressourcen** des MT-Anbieters und von den Möglichkeiten der Terminologieintegration.

# MT bis der Arzt kommt



## Die Herausforderung

Ein Pharmakonzern möchte die zentrale Übersetzung aller Dokumente auf **maschinelle Übersetzung + Post-Editing** umstellen. Zentrale Stakeholder sollen sukzessive Zugang zu diesem MT-Angebot bekommen. Zusätzlich sollen alle Nutzer des Unternehmens Zugriff auf einen **MT-Self-Service** bekommen. Ein passender Systemhersteller für mehr als 30 mit Kundendaten trainierten Engines wird gesucht.

## Unsere Lösung

Eine Vorauswahl von 16 **potenziellen Systemanbietern** wurde auf funktionale und technische Eignung untersucht. In die enge Auswahl kamen drei Anbieter, deren MT-Engines mit Kundendaten im Rahmen eines **Proof-of-Concepts** getestet wurden. Pro Anbieter wurden vier MT-Engines **trainiert** und **evaluiert**. Nach Abgleich aller Ergebnisse wurde das Siegersystem vom Kunden ausgewählt.



## Der Weg

Vor der Produktivnahme der MT-Engines wurden Prozesse für das Verfolgen der **MT-Qualität** etabliert. Relevante **Terminologieressourcen** werden laufend in kundenspezifischen MT-Profilen aufgearbeitet. In regelmäßigen Abstimmungen mit dem Systemhersteller werden **Optimierungspotenziale** abgestimmt und technische Voraussetzungen für die zentrale MT-Bereitstellung im Unternehmen geschaffen.

## Das Ergebnis

Das Unternehmen hat bereits 30 trainierte MT-Engines im Einsatz. Das Ziel ‚schnellere Übersetzung bei mehr Effizienz‘ wurde erreicht. Durch laufenden **Wissenstransfer** zwischen berns language consulting und dem Global Language Management Team ist der Kunde **mit allen relevanten MT-Aspekten vertraut**. Für technische Umsetzungen und im Falle von Ressourcenknappheit leistet blc dem Kunden weiterhin Support.



# Wie kommen Sie zu Ihrer eigenen MT-Lösung? Mit uns!

Die Orientierung in der stetig wachsenden MT-Anbieterlandschaft, Beurteilung und Extraktion von Trainingsressourcen, Training und Bewertung der Übersetzungsqualität – all das stellt Anwender vor neue Herausforderungen. Mit dem richtigen **Partner** werden diese problemlos bewältigt!

Im persönlichen Gespräch erörtern wir die **Umsetzung Ihrer individuellen maschinellen Übersetzungslösung**. Wir begleiten Sie auf dem Weg von der Anforderungsermittlung zur fertigen MT-Lösung bis in den Betrieb. Unsere MT-Expert:innen freuen sich auf ein Gespräch mit Ihnen!

Ihre Expert:innen für Maschinelle Übersetzung:



**Christian Eisold**

☎ 0176 / 10 09 72 48

✉ [c.eisold@berns-language-consulting.de](mailto:c.eisold@berns-language-consulting.de)



**Sophia Ackermann**

☎ 0151 / 21 45 87 62

✉ [s.ackermann@berns-language-consulting.de](mailto:s.ackermann@berns-language-consulting.de)

# Sie haben Fragen? Kontaktieren Sie uns!

**berns language consulting GmbH**

Blücherstraße 69  
40477 Düsseldorf

Zentrale +49 (0) 211 22 06 77 0

[www.berns-language-consulting.de](http://www.berns-language-consulting.de)

 [linkedin.com/company/berns-language-consulting](https://www.linkedin.com/company/berns-language-consulting)

**Besuchen Sie unsere Leistungsseite zu  
Maschinellem Übersetzung**

